Performance evaluation of popular *l1*-minimization algorithms in the context of Compressed Sensing

Bijeesh TV
Department of Computer Science and Engg, Sree Narayana Guru College of Engg & Technology, Kannur, India.
bijeeshtv@gmail.com

ABSTRACT

Compressed sensing (CS) is a data acquisition technique that is gaining popularity because of the fact that the reconstruction of the original signal is possible even if it was sampled at a sub-Nyquist rate. In contrast to the traditional sampling method, in CS we take a few measurements from the signal and the original signal can then be reconstructed from these measurements by using an optimization technique called *l1*-minimization. Computer engineers and mathematician have been equally fascinated by this latest trend in digital signal processing. In this work we perform an evaluation of different *l1*-minimization algorithms for their performance in reconstructing the signal in the context of CS. The algorithms that have been evaluated are PALM (Primal Augmented Lagrangian Multiplier method), DALM (Dual Augmented Lagrangian Multiplier method) and ISTA (Iterative Soft Thresholding Algorithm). The evaluation is done based on three parameters which are execution time, PSNR and RMSE.

Keywords: Compressed sensing (CS), sub-Nyquist, *l1*-minimization, PALM, DALM, ISTA.

1. INTRODUCTION

Compressed sensing has gained its popularity due to the fact that it provides a means to reconstruct the signal despite violating the much celebrated Shannon-Nyquist sampling theorem. The traditional method of taking samples from a signal is no longer used in compressed sensing. Instead we take only a few measurements from the signal using which the signal can be reconstructed. Care should be taken while taking these measurements because the efficiency of this process in turn affects outcome of the reconstruction algorithm. Though the crux of compressed sensing is the sub-Nyquist sampling, the signal should satisfy an essential condition for CS to work. The condition is that the signal should be sparse in the mathematical domain in which the measurements are taken. Though this seems to be a strict restriction on CS, since most of the natural signals like speech and image can be found to be sparse in any one mathematical domain, this problem can be easily overcome. So it is enough to transform the signal to the mathematical domain in which the signal is sparse. Usually the signals are sparse in the

Frequency domain and thus we use Discrete Cosine Transform (DCT) or Discrete Fourier Transform (DFT) to transform the signal from the time domain or spatial domain to the frequency domain. Discrete Wavelet Transform (DWT) can also be used for sparcifying the signal. The mathematics involved in CS is discussed in detail in the subsequent sections.

Out of the many important aspects of CS, the choice reconstruction method is a very crucial one. This paper aims at finding out the appropriate reconstruction algorithm to be employed under different circumstances. We evaluate three most commonly used reconstruction algorithms for their performance. The power of CS lies in the design of an efficient measurement matrix and the use of an efficient reconstruction algorithm. In this paper we attempt to answer the latter issue.

1.1 HISTORY OF CS

Compressed sensing was discovered by Emmanuel Candès in 2004. Candès was trying to clean up a highly corrupted version of the Shepp-Logan Phantom image. He performed a mathematical technique called *l1*-minimization on the noisy phantom image expecting it to slightly enhance the image. The outcome of this experiment was very surprising even to Candès himself as it gave nearly impossible results. The output obtained was very sharp and the image was perfect in every detail. This was a revelation to the entire signal processing community and mathematicians alike. Candès then discussed this result with a colleague Terrance Tao and both of them together formulated the basic theory behind this 'magic'. That was the beginning of an entire new field of mathematics called Compressed Sensing or Compressive Sensing (CS). Candès explains the mathematics behind this in [1] and [2]. D.L.Donoho discusses CS in detail in his work [3]. A lot of researchers have taken to CS since Candès introduced it to the scientific community. CS has been used in numerous applications by scientists, researchers and students alike. The mathematical approaches in CS were presented in the work of Jianhua Zhou, Siwang Zhou and Qiang Fan [4]. The major applications of CS over the years have been medical image processing, satellite image processing, speech processing, network power optimization, digital communication etc. Yi Zhong and Jiahou Huang present an improved Reconstruction Algorithm based on Compressed Sensing for power quality analysis in wireless sensor networks in [5].

Since its birth, CS has been applied on numerous areas of research. The first application where CS was used was MRI imaging. Unlike the traditional MRI imaging which took relatively longer time to obtain the image, CS could extract enough samples (measurements) from which a high quality MRI image can then be reproduced with the help of *l1*-minimization. This was just a beginning, what followed was a flurry of algorithms based on CS from all parts of the world. CS was applied on almost all signal processing tasks. Researchers from Rice

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

University later developed a hardware based on CS, a single pixel camera. With traditional digital camera, samples are first taken and then the unwanted samples are dropped during compression. This is an absolute wastage of time and processing power and CS fits in to overcome this limitation perfectly. Instead of sampling and then dropping samples, CS enables to take only those samples that are essential in reconstructing the original signal and thereby improving the performance in terms of processing time and processing power. The single pixel camera uses a low-cost, fast and sensitive optical detection technique to take the measurements. The measurements are taken based on a random basis matrix. And these measurements are found to be sufficient to reconstruct the image using *l1*-minimization. This single pixel camera saves storage space, processing time and the hardware cost when compared to the traditional many pixel cameras.

2. MATHEMATICS OF CS

The whole theory of Compressed Sensing is based on the famous matrix equation $y = Ax$. Any signal can be mathematically thought of as a vector. The underlying matrix problem is to find *x*, given *y* and *A*. The formulation of the whole problem is given in this section.

$$y = Ax \tag{1}$$

Equation (1) is interpreted in CS context as explained below.

y is the observed signal, *A* is the measurement matrix and *x* is the original signal. Let us assume that the original signal *x* is of size *N X 1*. As per CS theory, we need to take only lesser measurements than the Nyquist rate for a good reconstruction of the signal. This requires the observed signal *y* to be of size much lesser than that of the original signal. Assuming the size of *y* to be *M X 1*, we require that $M << N$ i.e. *M* should be very very less than *N*. This can be accomplished by properly designing the measurement matrix *A* with size *M X N*. Effectively; *M* measurements are taken from a signal of size *N* with the help of a measurement matrix *A*. A lot of research has gone into the designing of a good measurement matrix. The better the measurement matrix the better would be the signal after reconstruction.

The CS theory has two aspects: first the sensing aspect or the process of taking the measurements at a sub-Nyquist rate and the second is the reconstruction aspect or the process of generating the original signal from the measurements obtained from the sensing phase. The sensing was performed by using measurement matrix or sensing matrix. The reconstruction is done by performing a mathematical technique called *l1*-minimization. But as per the CS theory, *l1*-minimization gives a good reconstruction only if the original signal is a sparse one. There are many different ways to transform a natural signal from non-sparse form to sparse form.

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

Mathematical transformation like Cosine Transform, Fourier Transform, Wavelet Transform etc can be used to achieve this. These transforms are used to transform the signal from the time or space domain to other mathematical domains in which the signal has a sparse representation. Before the *l1*-minimization can be applied, we are required to perform the sparcification on the signal. This has to be incorporated in the formulation as well; resulting in the following formulation.

$$y = A\psi\alpha \tag{2}$$

Here ψ is the matrix that transforms the signal from one mathematical domain to another domain where the signal can be represented as a sparse signal called the transformation matrix or sparcifying matrix. The vector α is the vector of co-efficients resulting from applying ψ on *x*.

$$x = \psi\alpha \tag{3}$$

Since x is of size $N\ X\ 1$, α is obviously of the same size and ψ should therefore be of size $N\ X\ N$.

As mentioned earlier, CS can be understood as a combination of two sub-problems: first taking measurements from a signal and then reconstructing the signal from these measurements. The vector y is the output of the first sub-problem of taking measurements at a rate less than the Nyquist rate. This follows from the fact that y is of much smaller size than the vector x that corresponds to the original signal. Solving the second sub-problem is trickier than the first one and utilizes the power of *l1*-minimization. CS theory is based on the fact that the minimum *l1*-norm solution to an underdetermined system of linear equations is also the sparsest possible solution under quite general conditions. Mathematically, suppose there exists an unknown signal $x \in \boldsymbol{R}^{\boldsymbol{n}}$, a measured vector $y \in R^d$ (d<<n), and a measurement matrix $\mathrm{A} \in \mathrm{R}^{d \times n}$ such that A is a full rank matrix and if $y = Ax$, then x can be exactly recovered by computing the minimum *l1*-norm solution.

The standard *l1*-norm function is taken as the objective function to be minimized with respect to the primal variable x. The objective function is minimized with respect to the linear constraint $Ax = y$. The whole CS problem thus can be formulated as an optimization problem. The formulation of the problem is,

$$\min_x \|x\|_1 \quad ; \text{such that } Ax = y \tag{4}$$

Many algorithms have been used for solving this optimization problem since CS became popular among researchers. This paper aims at evaluating some of the *l1*-

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

minimization algorithms in the context of CS. The algorithms being evaluated are PALM (Primal Augmented Lagrangian Multiplier method), DALM (Dual Augmented Lagrangian Multiplier method) and OMP (Orthogonal Matching Pursuit). In the following section a brief description of these algorithms is given.

3. L1-MINIMIZATION ALGORITHMS

3.1.1 ISTA (ITERATIVE SOFT THRESHOLDING ALGORITHMS)

This algorithm utilizes an approach called Majorization-Minimization (MM) that has become quite popular over the years and is useful for numerous problems in signal processing: denoising, deconvolution, interpolation, super-resolution, declipping, etc [6]. Majorization-minimization is a technique in optimization theory which proceeds by choosing a function that majorizes (maximizes) the objective function. After majorization step the chosen function is minimized to obtain the solution. The complete derivation of ISTA can be found in [7]. The derivation utilizes many concepts of linear algebra and vector derivatives. ISTA is a combination of two techniques called Landweber algorithm and Soft thresholding and therefore is also known by the name Thresholded-Landweber algorithm. Landweber algorithm is an iterative method in which the iteration continues till the solution converges. Landweber iteration is followed by soft thresholding to obtain the final solution.

3.1.2 AUGMENTED LAGRANGIAN MULTIPLIER METHODS

Primal Augmented Lagrangian Method and Dual Augmented Lagrangian methods are both formulated in a similar manner with the addition of a new term called as the Lagrangian Multiplier. Lagrangian Multiplier methods have become popular in the area of convex programming. In these methods we eliminate the equality constraints and add a penalty term to the objective function. Our objective is to solve the system of equations $Ax = y$. But in practical situations, the solution to this is likely to introduce some error or in other words Ax may not be exactly equal to y. Therefore for practical cases we consider $Ax \approx y$, which can be rewritten using the equality as $Ax + r = y$; where r is the error or the residual term. The objective thus becomes minimizing the norm of x as well as the norm of the residual r. The two variations of Augmented Lagrangian Multiplier methods are Primal ALM (PALM) and Dual ALM (DALM). In DALM, the dual form of the Lagrangian of the objective function is minimized to obtain the solution. A detailed description along with the derivation of ALM methods can be found in [8] and [9]. An insight into PALM can also be found in [10].

4. EXPERIMENTS AND RESULTS

In this paper the experiments are aimed at evaluating the *l1*-minimization methods discussed in the previous section in the context of CS. The measurement matrix A, used in our experiments is orthogonal random matrix and DCT matrix is used as the transformation matrix ψ. The evaluation is done based on execution time, Root Mean Squared Error (RMSE) and Peak Signal to Noise Ratio (PSNR). A speech signal with 1000 samples is used as x and we attempt the reconstruction from as little as 50 measurements.

Figures 1,2 and 3 shows the original signal against the signal reconstructed from different number of measurements using Primal Augmented Lagrangian Multiplier method, Dual Augmented Lagrangian Multiplier method and Iterative Soft Thresholding Algorithm respectively.

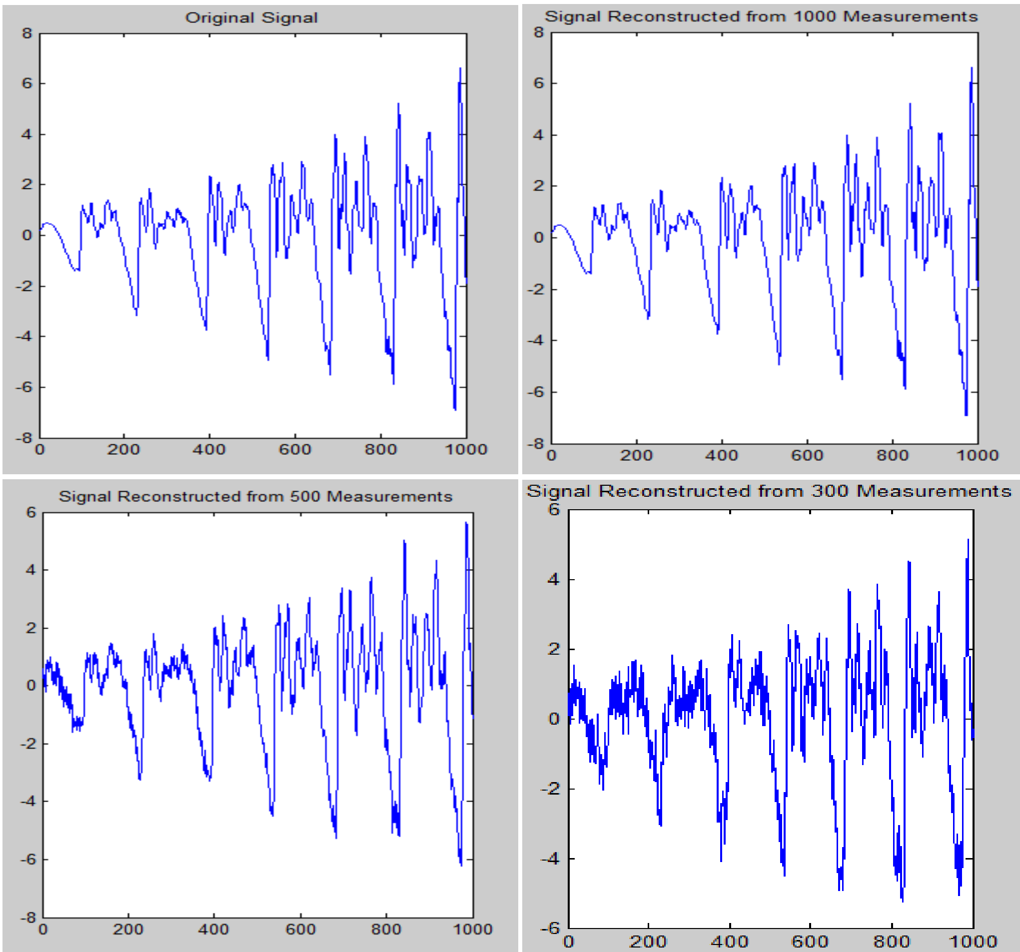


FIGURE 1: Signal reconstruction using PALM

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

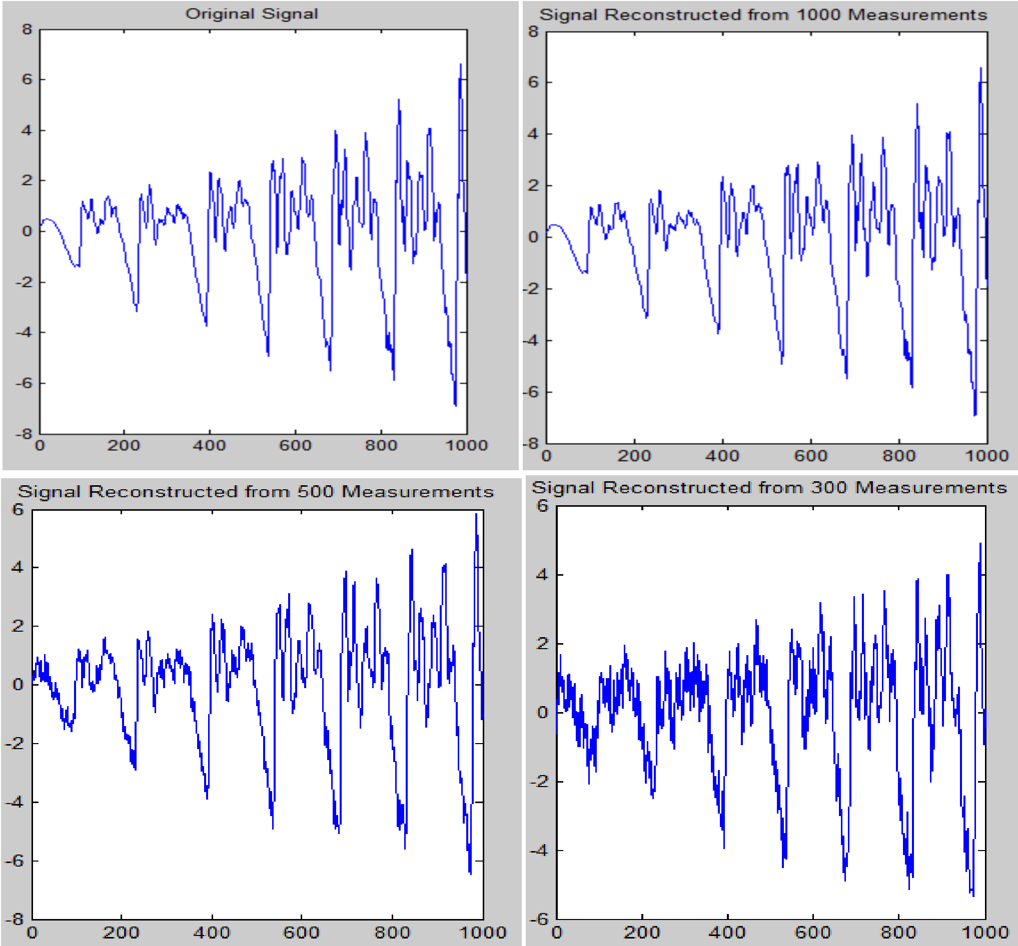
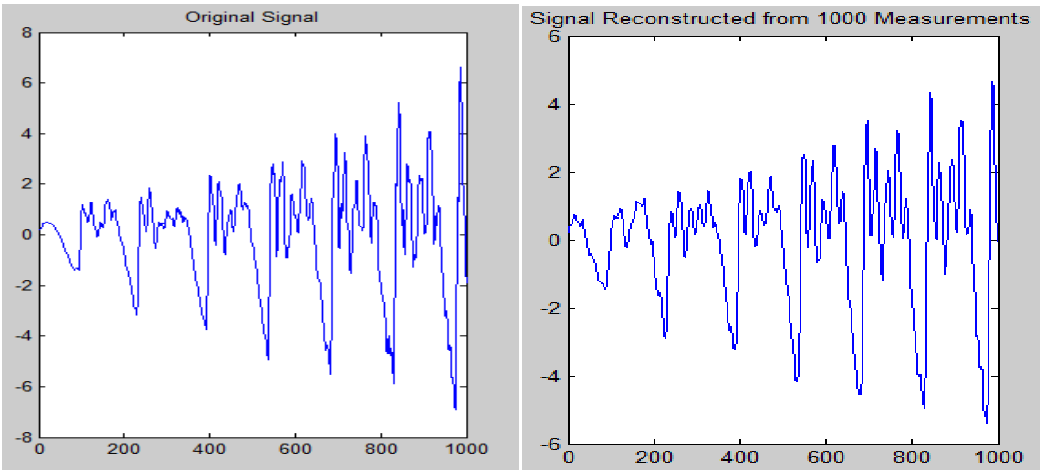


FIGURE 2: Signal reconstruction using DALM



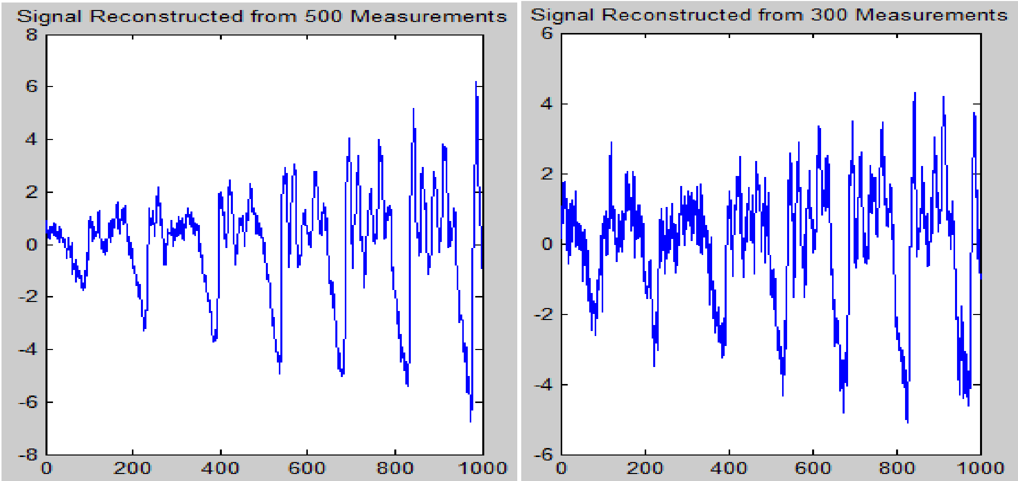


FIGURE 3: Signal reconstruction using ISTA

It is evident from the above figures that for a signal of 1000 samples, good reconstruction is possible with 300 measurements. For many applications, an acceptable level of reconstruction accuracy can be achieved from even lesser measurements. Table1 provides a detailed comparison of the three algorithms that were discussed in the previous sections.

TABLE 1.
Performance comparison of PALM, DALM and ISTA

No. of measurements	PALM			DALM			ISTA		
	Execution time(sec)	PSNR	RMSE	Execution time(sec)	PSNR	RMSE	Execution time (sec)	PSNR	RMSE
1000	1.331	116.189	0.001	1.137	62.958	0.005	0.593	24.516	0.412
900	3.892	46.953	0.031	0.964	49.195	0.024	0.586	48.943	0.026
800	3.923	42.334	0.053	0.868	40.933	0.062	0.549	41.768	0.056
700	4.056	35.686	0.114	0.835	34.563	0.129	0.487	37.757	0.089
600	3.639	30.722	0.202	0.634	32.108	0.172	0.459	32.141	0.171
500	3.543	26.934	0.312	0.552	27.218	0.301	0.413	28.157	0.271
400	3.195	23.219	0.478	0.496	22.881	0.497	0.393	23.403	0.468
300	3.379	20.066	0.688	0.371	20.442	0.659	0.297	18.978	0.779
200	2.854	16.561	1.029	0.209	17.178	0.959	0.254	16.036	1.093
100	3.346	12.735	1.599	0.169	13.201	1.516	0.348	12.512	1.641
50	2.909	10.169	2.149	0.126	10.459	2.079	0.192	11.037	1.944

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

The comparison results are better explained by the graphs presented in Figure4, 5 and 6. The graphs plot PSNR, RMSE and Execution time against the number of measurements taken.

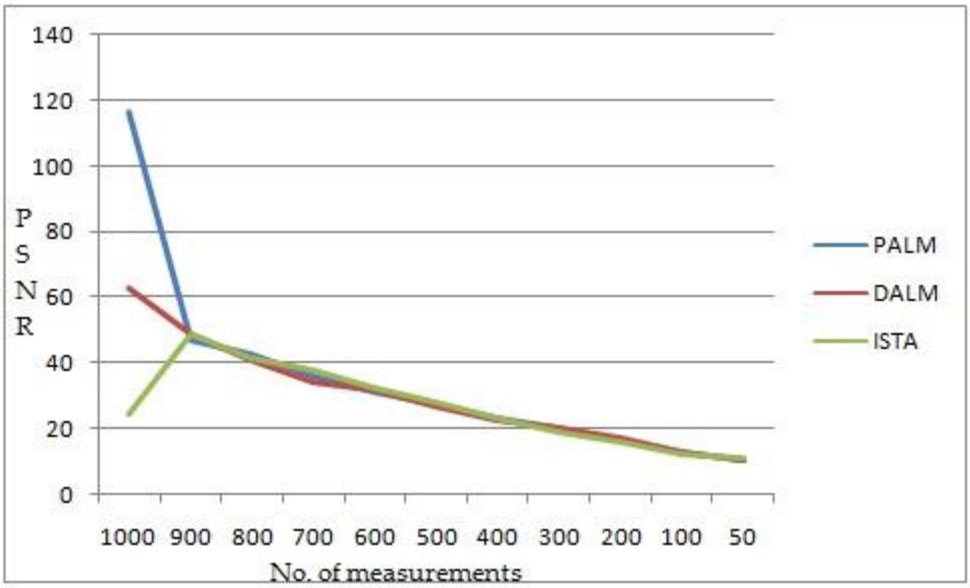


FIGURE 4. Comparison based on PSNR

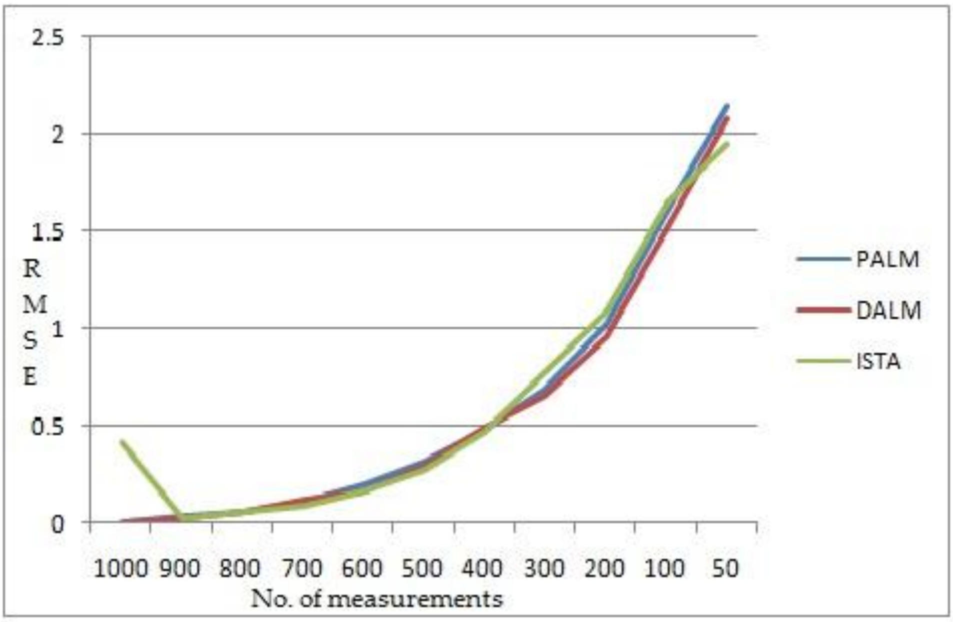


FIGURE 5. Comparison based on RMSE

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

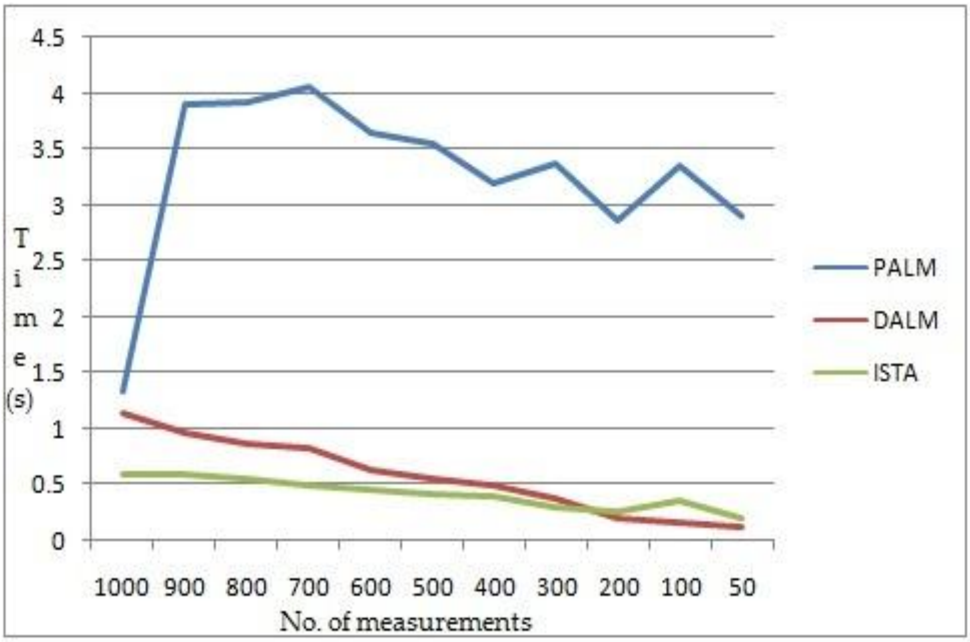


FIGURE 6. Comparison based on execution time

5. CONCLUSION

In this paper, three signal reconstruction algorithms are evaluated for the performance in the context of compressed sensing. Apart from performance evaluation this work emphasis the power of *l1*-minimization based reconstruction algorithms. The evaluation is done based on execution time, PSNR and RMSE. For PALM the execution time is found to be higher than that of the other two algorithms, but with respect to PSNR and RMSE, PALM produced results comparable with DALM and ISTA. Excellent reconstruction of the original signal is achieved with all the three algorithms. These results are obtained by using random matrix as the measurement matrix. By using a more carefully designed measurement matrix, CS would give much better results.

REFERENCES

[1] E. Candès, “Compressive Sampling,” Proc. Int'l Congress of Mathematicians, 2006.

[2] Candès, Emmanuel J. and Wakin, Michael B. ‘*An Introduction To Compressive Sampling [A sensing/sampling paradigm that goes against the common knowledge in data acquisition]*’. IEEE Signal Processing Magazine, 25 (2). pp. 21-30. ISSN 1053-5888. (2008).

[3] D. L. Donoho, ‘*Compressed Sensing, Manuscript*’, September 2004.

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

[4] Jianhua Zhou, Siwang Zhou, Qiang Fan.' *Mathematics Approaches in Compressed Sensing*'. TELKOMNIKA Indonesian Journal of Electrical Engineering *Vol 11 No 9, 2013.*

[5] Yi Zhong, Jiahou Huang.*'An Improved Reconstruction Algorithm Based on Compressed Sensing for Power Quality Analysis in Wireless Sensor Networks of Smart Grid*'. *TELKOMNIKA* Indonesian Journal of Electrical Engineering *Vol 12 No 8, 2014.*

[6] I. Daubechies, G. Teschke, and L. Vese. '*On some iterative concepts for image restoration*'. In P. W.Hawkes, editor, Advances in Imaging and Electron Physics, volume 150. Elsevier, 2008.

[7] I.W.Selesnik. '*Sparse Signal Restoration*' 2006

[8] Yin Zhang. 'User's *Guide for YALL1: Your ALgorithms for L1 Optimization'.* CAAM Technical Report TR09-17

[9] Allen Y. Yang, Arvind Ganesh, Zihan Zhou, S. Shankar Sastry, and Yi Ma. '*A Review of Fast l1-Minimization Algorithms for Robust Face Recognition*'. Proceedings of the International Conference on Image Processing in 2010.

[10] Premjith B, Sachin Kumar S, Akhil Manikkoth, Bijeesh T V, K P Soman. '*Insight into Primal Augmented Lagrangian Multilplier Method'.* International Journal of Computer Science and Information Technology & Security 2011.

[11] Compressive Sensing Resources, http://dsp.rice.edu/cs